

TOMASZ GACEK

Jagiellonian University in Kraków

tomasz.gacek@uj.edu.pl

## SOME REMARKS ON THE LANGUAGE OF MODERN TAJIK PUBLIC INSCRIPTIONS (PART I)

**Keywords:** Tajik language, public inscriptions, multilingual texts

### Abstract

The present article deals with the Tajik language used in modern public inscriptions (sign-boards, sign-posts, billboard advertisements, political banners, etc.) documented in about 400 photographs taken in Tajikistan by various individuals in recent years. Some sociolinguistic problems are discussed (especially in the case of multilingual inscriptions) as well as morphology, vocabulary, word-formation and syntax of the texts in question.

### 1. Introduction

The present article deals with the Tajik language used in modern public inscriptions, i.e. sign-boards, sign-posts, billboard advertisements, political banners, etc. The word “modern” is to be understood as “still in existence”, thus relatively old examples from the Soviet-era will be included, if they still can be seen today.

The inscriptions studied were collected as a result of research involving an analysis of about 400 photographs taken in Tajikistan by various individuals in recent years. The pictures were mostly accessed via online resources. As a result, around 270 inscriptions were chosen for further analysis. Most of the analyzed photographs were taken within the last two decades, although – in some cases – the date was not available so it may be difficult to date them precisely.

In some cases only a part of an inscription was possible to be read, due to either technical or compositional reasons.

A considerable number of the analyzed photographs were found on the Panoramio online service, which on February 2017 has closed, but the photos are expected to be available on Google Maps.

In the first part of the present series of publications languages used in the analyzed texts (apart from Tajik) will be discussed. An attempt at a classification of the inscriptions based on their content will follow. Then the language of the Tajik inscriptions will be analyzed, starting with phonetics, phonology and orthography.

Morphology, vocabulary, word-formation and syntax will be discussed in the following parts, to be followed by conclusions.

This analysis is – in particular – focused on forms and structures that are specifically Tajik, i.e. not used in the Persian of Iran.

As the inscriptions were collected from disparate sources and their availability may change over time, they are provided *in extenso* in the appendix to this article.

### 1.1. Languages of the analyzed texts

Not all of the analyzed inscriptions are written in Tajik, with examples of texts in other languages relatively numerous. While most of the analyzed inscriptions<sup>1</sup> are, indeed, either in Tajik or contain a variant in this language (227 examples, i.e. around 84% of the total 270). Russian is found in around 70 instances (26%). There are also more than 30 (11%) inscriptions in English, and four examples of Chinese. In the case of around 20 items (almost 7.5%), the language of a particular inscription (or one of its variants) was non-determinable. These are mostly certain very short texts (especially those consisting of a single word), which could equally well be Russian as Tajik. Only isolated examples of other languages are found, such as the bilingual Hebrew-Russian inscription at the entrance to the old synagogue in Dushanbe (now demolished). Together the analyzed texts may be considered as a multilingual mini-corpus of around 2000 words, with more than 60% of them in Tajik.<sup>2</sup>

A separate problem is the question of the potential use of Persian (i.e. the Persian of Iran, Fārsi) in the analyzed corpus. The relationship between Fārsi and Tajik is a complicated matter, and there is a dilemma as to whether to classify them as two dialects of one language or as two closely related languages.<sup>3</sup> In the case of written texts, the most obvious difference between standard Fārsi and Tajik is the script used, being – typically – Perso-Arabic and Cyrillic, respectively. Therefore, although it would be surprising to find a Fārsi text in Cyrillic script, Tajik – almost a hundred years after abandoning the Perso-Arabic script – is still occasionally written using this alphabet. Despite this, there was an attempt to reintroduce Perso-Arabic as the main writing system for the language in the 1990s (Landau 2010: 28). Thus, it is not possible to classify an inscription as being written in Fārsi on the basis of the script alone. Instead, lexical or grammatical elements specific to Fārsi

<sup>1</sup> Some analyzed texts are bi- or even tri-lingual, so the total number of the texts exceeds by far the number of photographs from which they are derived.

<sup>2</sup> Expanding the database of modern Tajik inscriptions may be a valuable enterprise itself, even if some Tajik language corpora already exist (Dovudov et al. 2011: 89ff.).

<sup>3</sup> On the relationship between Persian and Tajik in wider historical perspective see Rzehak (2001).

should be determined, that is those not found in Tajik, or clear indications that the intention was to write an inscription in Fārsi, e.g. to make it accessible to Iranian visitors. Having said that, it has to be admitted that Fārsi texts may be found in the analyzed corpus, although they are extremely rare.

It is worth noting that among the inscriptions in languages other than Tajik, it is only Russian that can be used independently, i.e. with no equivalent in another language. The evidence for this is provided in examples such as *БОТАНИЧЕСКИЙ САД* [6], *СОЛЯРКА* [13], *МЯСОКОНСЕРВНЫЙ КОМБИНАТ* [5]. Moreover, in a multilingual inscription with no Tajik version, without doubt there will be a Russian variant (mostly alongside English), e.g. *MERCURY | ИНТЕРНЕТ КАФЕ | ИНТЕРНЕТ | КСЕРОКОПИЯ | СКАНИРОВАНИЕ (...)* [7], *SHOP | РАМИР | МАГАЗИН | ПАМИР* [16]. In some cases, however, there are bilingual Russo-Chinese inscriptions [30], and even one example of a English-Chinese text has been found: 接待 | *Reception* [21].<sup>4</sup>

Finally, there are examples of inscriptions, the language of which cannot be determined. In practice, in most cases this means that they may be either in Tajik or Russian. These are mostly very short texts, in which syntactical features cannot be observed, while their vocabulary is common to both Tajik and Russian, e.g. *СТОМАТОЛОГИЯ* [31] and *АВТОДРОМ* [28].

## 1.2. Writing systems

As far as the writing systems used in the analyzed inscriptions are concerned, Cyrillic script is certainly the most universally used (both for Tajik and Russian). However, in the case of Tajik, there are a number of texts in Perso-Arabic script<sup>5</sup> (e.g. (...) آرامگاه [14] or برتر از دیدار روی دوستان [18] and at least one historical inscription containing a Tajik text written using the Latin orthography of the 1920–1930's, *МАКТАБИ МИЈОНАҶИ N 20 (...)* [4]. Apart from this rare example, Latin script is mostly used as the alphabet of English. Historically the Latin script was used extensively in the 1920–1930's, including in public inscriptions, for instance, the official emblems of the Tajik Autonomous Soviet Socialist Republic and subsequently the Tajik Soviet Socialist Republic, e.g. the state arms of Tajik ASSR: *Çumhūrijati sūsiolistiĵi şūraviji muxtori Toçikston*, accompanied by a variant in the Perso-Arabic script (Borĵian 1999). Nowadays, however, Tajik inscriptions in Latin script are extremely rare in the public sphere.

Chinese characters appear in a number of Chinese-language inscriptions, 中塔友谊车 [3], 接待 [21] (see above).

## 1.3. Classification of inscriptions

All the inscriptions were classified on the basis of their content:

<sup>4</sup> I would like to express my gratitude to prof. Ewa Zajdler for proofreading the Chinese examples.

<sup>5</sup> See the discussion on Fārsi vs. Tajik forms above.

Inscription category	No. of occurrences
signboards	140
political posters	37
advertisements	30
festival greetings	21
memorial inscriptions	12
signposts	8
announcements	6
other	25

These categories may overlap, as, for instance, festival greetings are often, at the same time, political posters, e.g. *Наврӯз – ҳумоюнномаи | худдиносиву худогоҳии | миллати тоҷик аст! | Эмомалӣ Раҳмон* [27].

It is particularly interesting to consider the results of this classification in junction with the data on the languages used in particular inscriptions. As we have seen, most of the analyzed inscriptions are signboards and the Tajik language is used in about 85% of these (nearly 60% are monolingual), while a Russian variant is present in around 25%, with, significantly, around 15% in English.

In the case of advertisements Tajik is prevalent as well (80%), however, Russian has a much greater presence (over 53%). Apart from that only English is attested (around 10%).

If we look at political posters the domination of Tajik is overwhelming (95%), with only 2 isolated examples in Russian.

An extreme example is provided by festival greetings, where no language other than Tajik is attested.

On the other hand, memorial inscriptions – owing to the presence of numerous World War II monuments in the Tajik landscape – show a stronger presence of Russian (in 4 of 12 inscriptions), even if Tajik is still the most important medium in this category too (10 examples). English and the Persian of Iran are attested in this category as well. An interesting example is a four-language inscription *Панҷакентӣ қадим асрҳои V–VIII | پنچکنت قدیم قرنهای ۵–۸ م | Древний Пенджакент V–VIII века | Ancient Panjkent 5–8 centuries* [22]. The context (an inscription for tourists, possibly from Iran, as can be seen from the difference *asr* – *qarn*) and the presence of a definitively Tajik Cyrillic variant suggests this is a rare example of a Fārsi text in the analyzed corpus.

As far as signposts and similar geographical or topographical inscriptions are concerned, most consist only of proper names and, as proper names are – at least to some extent – language independent, their language is often non-determinable. Wherever it is possible to indicate a particular language, it is mostly Tajik or in one case Russian. Various everyday announcements form only an extremely small subsection of the analyzed inscriptions (6 examples). All are written solely in Tajik.

Finally, of the 26 remaining (i.e. impossible to be classified into one of the previous categories) inscriptions, Tajik is used in 21 (18 have no other language variant). Interestingly, within this group there are 3 examples of Chinese, with Russian having 4 occurrences and English 2.

## 2. Language of Tajik inscriptions

### 2.1. Phonetics, phonology and orthography

Of course, inscriptions are not usually the first choice for a researcher of the phonetics and phonology of a living language, with native-speakers available and innumerable recordings and broadcasts accessible. Nevertheless, in the case of Tajik, where various variants of the language coexist, it may be of interest to determine which pronunciations are favoured by the orthography of the inscriptions. Thus, a few comments will be made on this topic.

How profound the influence of Russian was, and still is, can be shown not only by the number of loanwords but also by the influence it had on Tajik phonetics and phonology. Forms like *сканер* [26] and *плазма* [19] suggest that Tajik, contrary to the Persian of Iran, developed a kind of acceptance of initial consonant clusters. Also the analysis of audio recordings proves that this is certainly not only an orthographic convention (Gacek 2012: 356–357).

Some phenomena are particularly interesting if we compare the actual forms with information on the reforms of Tajik orthography. As we have already noted, Tajik has been written in three different scripts: (Perso)-Arabic, Latin and Cyrillic. After the fall of the Soviet Union, switching to Latin or Arabic were both considered (Landau 2010: 28). However, finally Tajik remained written in the Cyrillic script, although the orthography underwent minor changes to better reflect Tajik phonology and to de-Russianize it in a sense (Landau 2010: 28). In the analyzed corpus we find examples following the pre-reform orthography (e.g. *концертӣ* [24]) and also those that have accommodated the new rules (e.g. *коррупсия* [20]).

Tajik Cyrillic orthography depends on the use of diacritical marks in the notation of sounds unknown in Russian. Omitting these marks is – in general – not acceptable, however, examples of Tajik inscriptions written without diacritical marks have been found, e.g. *ФУРУШГОҶИ МАРКАЗИ | ЦУМ | ШИРКАТИ САҶОМИ* [1]. Apart from these, there is a considerable number of examples of the diacritics being used somewhat inconsistently, e.g. *МАҶОЗАИ ХУРОКВОРИ* (instead of the expected \**МАҶОЗАИ ХУРОКВОРИЙ*) [15], *Донишгоҳи давлатии Кӯлоб ба номи Рудаки* (instead of \**Рӯдакӣ*) [9] c.f. *БОҶИ УСТОД РҶДАКӢ* [17].

From the point of view of pronunciation, forms like *интернатсионалист* [25] ‘an internationalist’, *конститутсия* [23] ‘constitution’, *федератсия* [32] ‘federation’, *милитсия* [29] ‘militia’, etc. betray the fact that despite removing the letter *ч* from the Tajik orthography, at least some speakers of the language may be using the affricate /ts/ (written with two letters, *тс*). In fact, even modern Tajik dictionaries

include the form *милиция* beside the de-Russianized *милиса*.<sup>6</sup> Moreover, there are also examples of a letter now no longer used in Tajik orthography, the letter *ц*, e.g. *кино-концертӣ* [24] ‘cinema and concert [hall]’. On the other hand, other inscriptions render the original Russian affricate with the letter *с* alone, indicating a spirantic articulation, e.g. *коррупция* [20] ‘corruption’. As far as the actual pronunciation is concerned, this seems to be a feature of an individual’s own pronunciation. Many speakers of Tajik pronounce [s] rather than [ts]. Nevertheless, the affricate is clearly audible in the speech of others (Gacek 2012: 358).

Another interesting phenomenon related to the Tajik-Russian linguistic interaction is the problem in such forms as *автовокзал* [10, 12] ‘bus station’. The form in question is stressed, both in Russian and Tajik, on the last syllable (this stable accent in Tajik coincides with Russian in this instance). Thus, in Russian, because of a well-known process called *akanye*, the vowels of the second and third syllables should change from *-o-* to *-v-* (possibly the vowel in the second syllable could be further reduced). However, the modern Tajik orthography of the word shows no trace of this phenomenon, as it relies on the original (i.e. Russian) written form. Nevertheless, research into the actual pronunciation of Russian loanwords potentially undergoing this process reveals that some speakers follow the standard Russian pronunciation, despite its incompatibility with the written form (Gacek 2012: 355). Moreover, the form in question reveals other features of specifically Russian pronunciation, observable in Tajik, as well. Namely, [v] is pronounced as [f] before [t] and [k] is voiced to [g] before [z].<sup>7</sup>

There are also examples of the influence of the modern Persian of Iran (Fārsi) on the Tajik ethnolect. Most of the native forms in the analyzed corpus belong to the common Fārsi-Darī-Tajik stock, however, in isolated cases influences can be seen between the closely related idioms, for example, the form *бонк* [2, 8, 11 etc.]. Such a form is also attested in modern dictionaries.<sup>8</sup> Older dictionaries also mention the form *банк*<sup>9</sup> as well as *бонк*. While *банк* corresponds directly to the Russian (a possible donor language), *бонк* may potentially be explained only as a learned borrowing from Fārsi *بانک* /bānk/, learned in the sense that it complies with the rule that the Tajik /o/ corresponds to the Fārsi /ā/.

## References

- Bertel’s Ye.E. et al. 1954. *Tadzhiksko-Russkij Slovar’*. Moskva.  
 Borjian H. 1999. FLAGS ii. of Tajikistan. – *Encyclopædia Iranica*. [<http://www.iranicaonline.org/articles/flags-iii>]; last updated 2012; accessed on 2017-01-23; available in print: vol. 10/1]: 32–35.

<sup>6</sup> Saymiddinov et al. (2006: 116).

<sup>7</sup> The author would like to express his gratitude to the anonymous reviewer of the present article for drawing his attention to this point.

<sup>8</sup> Bertels et al. (1954: 45).

<sup>9</sup> Bertels et al. (1954: 79, where it is marked as a conversational, spoken form).

- Dovudov G. et al. 2011. Building a 50m corpus of Tajik language. – *RASLAN [Recent advances in Slavonic natural language processing]*; <https://nlp.fi.muni.cz/raslan/raslan11.pdf>]: 89–95.
- Gacek T. 2012. Some remarks on the pronunciation of Russian loanwords in Tajik. – *SLing* 129 [supplementum]: 353–361.
- Landau J.M. 2010. Alphabet reform in the six independent ex-Soviet Muslim republics – *JRAS* [Series 3] 20.1: 25–32.
- Rzehak L. 2001. *Vom Persischen zum Tadschikischen. Sprachliches Handeln und Sprachplanung in Transoxanien zwischen Tradition, Moderne und Sowjetmacht (1900–1956)*. Wiesbaden.
- Saymiddinov D. et al. 2006. *Farhangi Tojiki ba Rusi*. Dušanbe.

## Appendix

### List of inscriptions

The author may be mentioned by name (wherever provided) or a nickname. Where neither of these is to be found, the owner of the website or the website itself is given.

Ideally, the date of a photograph refers to the day when it was taken. Where this is not explicitly mentioned, other options are used, in the following order: publication date, upload date, date extracted from picture file, access date.

1. КРБИСМАН 2008-6-28a Dushanbe ФУРУШГОҶИ МАРКАЗИ | ЦУМ | ШИРКАТИ САҶОМИ [at:] <http://static.panoramio.com/photos/large/11645942.jpg>
2. Шахриёр Ёров 2013-6-20a Dushanbe [Тоҷик] Содирот Бонк | Tojik Sodiroṭ B[onk] [at:] <http://www.panoramio.com/photo/96821497>
3. lqf 2009-3-9 Dushanbe 中塔友谊车 | Автобуси дӯстии Тоҷикистону Хитой [at:] <http://www.panoramio.com/photo/26642441>
4. Burhon Tohirov 2014-05-13 Dushanbe МАКТАБИ МИҶОНАҶИ N 20 | СРЕДНЯЯ ШКОЛА N 20 | 1938 Г. [at:] <http://www.panoramio.com/photo/106914743>
5. onipkosergey 2011-2-4 Norak БОЗОРИ ШАҲРИ НОРАК [at:] <http://www.panoramio.com/photo/47548171>
6. DXT 1 2007-1-23 Shughnon БОТАНИЧЕСКИЙ САД [at:] <http://www.panoramio.com/photo/542426>
7. DURUN 2011-7-28 Khorugh MERCURY | ИНТЕРНЕТ КАФЕ | INTERNET CAFÉ | ИНТЕРНЕТ | КСЕРОКОПИЯ | СКАНИРОВАНИЕ | ЛАМИНАЦИЯ | (...) [at:] <http://www.panoramio.com/photo/56453879>
8. Dick van Santen 2012-7-11 Khorugh Аввалин Бонки Молиявии Хурд | The First Micro-Finance Bank [at:] <http://www.panoramio.com/photo/75189232>
9. sulaymon shobek 2007-7-6 Kulob Донишгоҳи давлатии Кӯлоб ба номи Рудаки (sic) [at:] <http://www.panoramio.com/photo/4443240>
10. Rustam Khalikov 2007-9-1 Qurghonteppa АВТОВОКЗАЛИ ҚУРҶОНТЕППА [at:] <http://www.panoramio.com/photo/11568437>
11. kasp 2008-7-26c Vahdat ОРИЁНБОНК [at:] <http://www.panoramio.com/photo/58297871>
12. fara12988 2012-5-25 Farkhor АВТОВОКЗАЛ ФАРХОР [at:] <http://www.panoramio.com/photo/69064271>
13. farflungistan 2012-9-4 Istaravshan солярка [at:] <https://www.flickr.com/photos/farflungistan/13923960753/>



14. shahinbedawi 2012-4-21 Kulob ۷۸۶-۷۱۴ | میر سید علی ہمدانی | عالم کبیر و عارف شہیر [at:] <https://www.flickr.com/photos/shahinbedawi/7105309993>
15. Joao Leitao JoaoleitaoTRAVEL 2012-6-8 Khorugh МАҶОЗАИ ХУРОКВОРИ (sic) [at:] <https://www.flickr.com/photos/joaoleitao/7995964418/>
16. Joao Leitao JoaoleitaoTRAVEL 2012-6-8 Khorugh shop | МАҶАЗИИ | ПАМИР | ПАМИР [at:] <https://www.flickr.com/photos/joaoleitao/7995957046/>
17. Maris Teteris 2010-4-10 Dushanbe БОҶИ УСТОД РӮДАКӢ [at:] <http://www.panoramio.com/photo/50270542>
18. Don Croner (?) 2012-12-25 Panj Rud برتر از دیدار روی دوستان [at:] <http://www.doncronerblog.com/2012/12/tajikistan-more-on-rudaki.html>
19. Menu.tj 2014-10-11 Dushanbe PLAZMA | НУҚТАИ ФУРӯШИ | ТЕЛЕФОНҲО | “ПЛАЗ-МА” [at:] <http://menu.tj/places/calon-svyazi-plazma/>
20. Ozodagon 2014-9-5 Tajikistan КОРРУПЦИЯ | Муқовимати ту | аҳамияти калон дорад! [at:] <http://www.ozodagon.com/17944-peshniodi-honanda-misli-kdak-dar-shabakaoi-itimo-girya-nakuned.html>
21. Maksim 2013-4-27 Dushanbe 接待 | Reception [at:] <http://yablor.ru/blogs/dushanbe-moskva/2847399>
22. 2014-10-25 Panjikent Панҷакенти қадим | асрҳои V-VIII || ۸-۵ قرضای قدیم | پنجکنت قدیم || Древний Пенджакент | V-VIII века || Ancient Panjikent | 5-8 centuries [at:] <http://www.azianatravel.com/en/ancient-penjikent>
23. Arnis Balcus 2010-10-10a Panjikent Конституция | кафолати ҳуқуқу | озодиҳои тоҷат [at:] <http://balcus.blogspot.com/2010/10/penjikent.html>
24. Arnis Balcus 2010-10-10 Panjikent Деваитиш | ЗАЛИ | КИНО-КОНЦЕРТИИ | ТО-БИСТОНА [at:] <http://balcus.blogspot.com/2010/10/penjikent.html>
25. Arnis Balcus 2010-10-10 Panjikent МО ИНТЕРНАЦИОНАЛИСТОНЕМ [at:] <http://balcus.blogspot.com/2010/10/penjikent.html>
26. Ozodagon 2014-3-19 Dushanbe ҲИЗМАТРАСОНИҲОИ ИНТЕРНЕТӢ || ИНТЕРНЕТ | НУСХАБАРДӢ | ЧОПКУНӢ | СКАНЕР | САБТ КАРДАНИ CD BA DVD [at:] <http://ozodagon.com/15406-orat-i-muoironi-toik-az-vatan-sar-meshavad-aks.html>
27. Ozodagon 2013-3-20 Dushanbe Наврӯз – ҳумоюнномаи | худшиносиву ҳудодоҳии | миллати тоҷик аст! | Эмомалӣ Раҳмон [at:] <http://www.ozodagon.com/9197-akso-shioroi-navrz-dar-kchavu-paskchaoi-dushanbe.html>
28. Flickr 2013-7-5 Kurghonteppa АВТОДРОМ [at:] <https://www.flickr.com/photos/79083427@N02/9215076884/>
29. Radyo-i Ozodi 2014-11-20 Kulob МИЛИТСИЯ [at:] <http://www.ozodi.org/content/local-police-officer-accused-of-religious-extremism-released-after--6-months/26701500.html>
30. Radyo-i Ozodi 2014-5-8 (...) [in Chinese] | Образцовый центр китайской сельскохозяйственной | техники и технологии | (...) [in Chinese] | Компания семенной промышленности «Henan Jingyan Silverseed Co LTD» КНР [at:] <http://www.ozodi.org/content/chinese-farmers-rent-100-hectars-lands-in-qumsangir-tajikistan/25377823.html>
31. Tajmedun 2015-1-16 Tajikistan СТОМАТОЛОГИЯ [at:] <http://tajmedun.tj/qumatib/fakul/stom/stomatologija.jpg>
32. Tojnews 2014-12-6 Tajikistan МАРОСИМИ ҶОИЗАСУПӢИИ | ФЕДЕРАТСИЯИ ФУТБОЛИ ТОҶИКИСТОН 2014 || Рӯзи 5 декабри соли 2014 соати 14:00 | дар Маҷмааи давлатии “Қоҳи Ваҳдат” [at:] <http://www.tojnews.org/tj/news/bekhtarinkhoi-futboli-tochikiston-dar-soli-2014-aks>